GPBR Exercise 4

Name: Soi Zhi Wen

Matriculation Number: 22-132-245

1. $graph\ density, d = \frac{2m}{n(n-1)} = \frac{2(17)}{9(9-1)} = 0.4722$

   $\boldsymbol{A = \{1, 2, 3, 4\}}$ $\boldsymbol{B = \{5, 6, 7, 8, 9\}}$

   $c(C_A, C_B) = 7$

   $$\delta_{int}(C_A) = \frac{8}{12} \approx 0.6667$$

   $$\delta_{int}(C_B) = \frac{12}{20} = 0.6$$

   $$\delta_{int}(g|C_A, C_B) = \frac{0.6667 + 0.6}{2} \approx 0.6333$$

   $$\delta_{ext}(g|C_A, C_B) = \frac{20}{40} = 0.5$$

   $\boldsymbol{C = \{1, 2, 3, 5\}}$ $\boldsymbol{D = \{4, 6, 7, 8, 9\}}$

   $c(C_C, C_D) = 3$

   $$\delta_{int}(C_C) = \frac{12}{12} = 1$$

   $$\delta_{int}(C_D) = \frac{16}{20} = 0.8$$

   $$\delta_{int}(g|C_C, C_D) = \frac{1 + 0.8}{2} = 0.9$$

   $$\delta_{ext}(g|C_C, C_D) = \frac{28}{40} = 0.7$$

   $\boldsymbol{E = \{1, 2, 3, 6, 8\}}$ $\boldsymbol{F = \{4, 5, 7, 9\}}$

   $c(C_E, C_F) = 10$

$$\delta_{int}(C_E) = \frac{8}{20} = 0.4$$

$$\delta_{int}(C_F) = \frac{6}{12} = 0.5$$

$$\delta_{int}(g|C_E, C_F) = \frac{0.4 + 0.5}{2} = 0.45$$

$$\delta_{ext}(g|C_E, C_F) = \frac{14}{40} = 0.35$$

Based on the computed metrics, we can argue that the clustering method that should be favored is the second one, $C = \{1,2,3,5\}$ $D = \{4,6,7,8,9\}$. This is because it has the highest intra-cluster density of all three clusterings, 0.9, indicating that nodes within each cluster are more densely connected. Additionally, the cut size between the two clusters, 3, is the smallest among the three clusterings, indicating that there are fewer edges connecting nodes from different clusters. This leads to a higher inter-cluster density, 0.7, than the other two clusterings, indicating that nodes in different clusters are less connected than nodes within the same cluster.

2. $I_{v_1} = 1$
   $E_{v_1} = 1$
   $D_{v_1} = E_{v_1} - I_{v_1} = 0$

   $I_{v_2} = 1$
   $E_{v_2} = 2$
   $D_{v_2} = E_{v_2} - I_{v_2} = 1$

   $I_{v_4} = 0$
   $E_{v_4} = 2$
   $D_{v_4} = E_{v_4} - I_{v_4} = 2$

   $I_{v_5} = 0$
   $E_{v_5} = 2$
   $D_{v_5} = E_{v_5} - I_{v_5} = 2$

   $I_{v_6} = 1$
   $E_{v_6} = 2$
   $D_{v_6} = E_{v_6} - I_{v_6} = 1$

   $I_{v_7} = 1$
   $E_{v_7} = 1$
   $D_{v_7} = E_{v_7} - I_{v_7} = 0$

$R_{1,4} = 0 + 2 - 2 = 0$
$No\ reduction\ of\ the\ cut\ size\ c(C_A, C_B)$

$R_{1,6} = 0 + 1 = 1$
$Reduction\ of\ the\ cut\ size\ c(C_A, C_B)\ from\ 5\ to\ 4$

$R_{1,7} = 0 + 0 = 0$
$No\ reduction\ of\ the\ cut\ size\ c(C_A, C_B)$

$R_{2,4} = 1 + 2 - 2 = 1$
$Reduction\ of\ the\ cut\ size\ c(C_A, C_B)\ from\ 5\ to\ 4$

$R_{2,6} = 1 + 1 - 2 = 0$
$No\ reduction\ of\ the\ cut\ size\ c(C_A, C_B)$

$R_{2,7} = 1 + 0 = 1$
$Reduction\ of\ the\ cut\ size\ c(C_A, C_B)\ from\ 5\ to\ 4$

$R_{5,4} = 2 + 2 = 4$
$Reduction\ of\ the\ cut\ size\ c(C_A, C_B)\ from\ 5\ to\ 1$

$R_{5,6} = 2 + 1 - 2 = 1$
$Reduction\ of\ the\ cut\ size\ c(C_A, C_B)\ from\ 5\ to\ 4$

$R_{5,7} = 2 + 0 - 2 = 0$
$No\ reduction\ of\ the\ cut\ size\ c(C_A, C_B)$

As a result, $R_{5,4}$ is the best solution which reduce the cut size (C_A, C_B) from 5 to 1.

3.

$\mathbf{L}(g) =$

```
[[ 4 -1 -1 -1 -1  0  0  0  0  0  0  0  0  0]
 [-1  3 -1 -1  0  0  0  0  0  0  0  0  0  0]
 [-1 -1  4  0 -1  0 -1  0  0  0  0  0  0  0]
 [-1 -1  0  3 -1  0  0  0  0  0  0  0  0  0]
 [-1  0 -1 -1  4  0  0  0 -1  0  0  0  0  0]
 [ 0  0  0  0  0  3 -1 -1 -1  0  0  0  0  0]
 [ 0  0 -1  0  0 -1  4 -1 -1  0  0  0  0  0]
 [ 0  0  0  0  0 -1 -1  3 -1  0  0  0  0  0]
 [ 0  0  0  0 -1 -1 -1 -1  5  0 -1  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  3 -1  0 -1 -1]
 [ 0  0  0  0  0  0  0  0 -1 -1  5 -1 -1 -1]
 [ 0  0  0  0  0  0  0  0  0  0 -1  2 -1  0]
 [ 0  0  0  0  0  0  0  0  0 -1 -1 -1  3  0]
 [ 0  0  0  0  0  0  0  0  0 -1 -1  0  0  2]]
```

$Fiedler\ vector, u_2 =$
$[-0.27195533, -0.2840522, -0.24299804, -0.2795814, -0.22602178,$
$-0.11372019, -0.14066911, -0.11372019, -0.06370238, 0.3619697,$
$0.2885418, 0.3619697, 0.3619697, 0.3619697]$

First partition:
$A = \{1,2,3,4,5,6,7,8,9\}$
$B = \{10,11,12,13,14\}$

$Fiedler\ vector\ of\ A, u_{2_A} =$
$[-0.33192355, -0.37781228, -0.17818838, -0.37781228, -0.17818837,$
$0.4310001, 0.29096233, 0.43100009, 0.29096234]$

Second partition:
$A_1 = \{1,2,3,4,5\}$
$A_2 = \{6,7,8,9\}$

Three clusters obtained:
$B = \{10,11,12,13,14\}$
$A_1 = \{1,2,3,4,5\}$
$A_2 = \{6,7,8,9\}$